Basics of Deep Learning And Neural Networks

Learning Outcomes

By the end of this topic, you will have achieved the following learning outcomes:

- Have a basic understanding of deep learning and neural networks.
- Have an understanding of how neural networks work.
- Have an understanding of the strengths and weaknesses of deep learning.
- Be able to create a neural network to solve a supervised learning problem.

"If you want to understand people, especially your customers, then you have to be able to possess a strong capability to analyze text." Paul Hofmann

Reading

What is Deep Learning?

Deep learning is a subset of Machine Learning that uses Neural Networks to simulate human-like decision making. ***Neural networks*** are models trained through learning from experience that simulates human-like decision making, and are used to effectively achieve better accuracy in predictions.

Deep learning can be supervised, unsupervised or semi-supervised, such that deep learning models can work with or without labelled data.

Examples of deep learning use cases are in cancer research where researchers are using deep learning models to automatically detect cancer cells. Teams use advanced microscopes that yield a high-dimensional data set which in turn is used to train a deep learning application/computer program to accurately identify cancer cells.

In other cases, deep learning models are used to help to improve worker safety around heavy machinery by automatically detecting when people or objects are within an unsafe distance of machines.

Why is Deep Learning useful?

- One of the reasons why deep learning is useful is because it addresses the shortcomings of traditional machine learning algorithms which may not be able to effectively handle high dimensional data. This would be data that has a large number of inputs and outputs. The handling of such data is usually complex and resource-intensive.
- Deep learning also automatically focuses on the right features during training. However, this is different when it comes to working with traditional machine learning algorithms. One would need to perform features selection in order to optimize the models.
- Deep learning also helps achieve higher accuracies as the magnitude of data increases, thus making it important for cases where safety-critical applications like in medicine and in self driving cars. In some cases, deep learning outperforms humans in specialized tasks i.e. classifying objects in images.

Applications of Deep Learning

Aside from the applications that we have briefly highlighted above, there are many other applications of deep learning in industry. Some of these applications include:

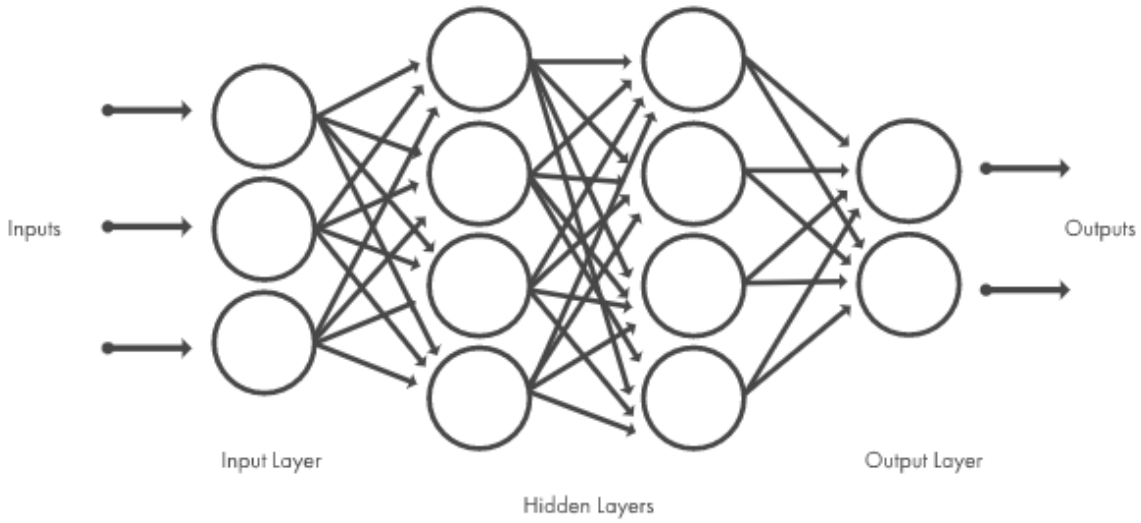
- Computer Vision
- Speech Recognition
- Automatic Machine Translation
- Virtual Personal Assistants
- Automated Self Driven Cars
- Facial Recognition and Automatic Tagging
- Chatbots

How Deep Learning Works

Deep learning models use **Artificial Neural Networks** (also called neural networks in short), which are adaptive systems that learn by using interconnected nodes or neurons in a layered structure that resembles a human brain.

Thus *neural networks* learn from data and as such, can be trained to recognize patterns, classify data, and forecast future events. The term "deep" refers to the number of hidden layers in a neural network.

The diagram below shows the architecture of a neural network.



[Image Source]

A neural network consists of several components. It consists of an *input layer*, one or more *hidden layers* and *an output layer*. In each layer, there are several nodes, or neurons, with each layer using the output of the previous layer as its input, so neurons interconnect the different layers. Each neuron typically has weights that are adjusted during the learning process, and as the weight decreases or increases, it changes the strength of the signal of that neuron.

Below is a breakdown of this process taken by the neural network to make predictions:

- **Input layer:**
  - This is the first layer of our neural network that consumes our data.
- **Hidden layer:**
  - Stacked layers between the input and output layer form a deep network. Each of the hidden layers is composed of neurons. In between the neurons there exists weights and biases which determine the values of those neurons.
  - There also exists another component referred to as *Activation Functions* for each layer, which transform the inputs of these neurons to a value between some lower limit and upper limit determined by these functions. Examples of these activation functions include *linear* or *identity*, *unit* or *binary step*, *sigmoid* or *logistic*, *tanh*, *reLU* and *softmax*. These activation functions determine whether a neuron would be activated, in order to be considered as inputs for the next neurons in the next layer.
  - During the training process and within the hidden layers, neural networks learn from labelled data overtime in order to make accurate predictions. They would do this via a process called *forward* and *backward*

propagation. You can watch the following video [Link] to get a visual sense of how this process happens.

- **Output layer:**
    - This is the last layer of our neural network that would produce the desired outputs/outcome.

There are no. of variants of artificial neural networks dedicated to solving specific problems i.e. Convolution Neural Networks are commonly used for image classification problems and Recurrent Neural Networks that are used to solve sequence problems. These variants will be covered in future sessions.

As this session will focus on creating basic neural networks, we will use the following steps to achieve this:

1. **Step 1: Data Importation**
2. **Step 2: Data Exploration**
3. **Step 3: Data Preparation**
4. **Step 4: Data Modeling**
    - Splitting our Dataset
    - Instantiating our Model
    - Adding Input, Hidden and Output Layers
    - Compiling our Model
    - Training our Model
    - Evaluating our Model
5. **Step 5: Summary of Findings**
6. **Step 6: Recommendations**

Limitations of Deep Learning

- **Complexity**
    - Unlike most machine learning algorithms, deep learning neural networks can be quite complex, especially when the no. of layers or neurons increase within a neural network.
- **Explainability**
    - If interpretability/ explainability is a success criterion than performance, then using other machine learning algorithms that may achieve similar or close to similar results, may be the best option.
- **Needs lots of Data**
    - To get the most out of deep learning models, there needs to be large amounts of data for training. Hence, deep learning models work well with high dimensional data.
- **Resource Intensive**
    - In order to fully effectively learn from its network, deep learning models require large computational power. This comes at a cost as deep learning

at a scale requires specialized processors i.e. Graphics Processing Unit (GPUs).

Deep Learning Tools

- **Keras**
  - This is an open-source neural network library written in python.
  - It provides industry-standard tested code for neural network model creation.
- **Pandas**
  - Python library for data manipulation and analysis.
- **Scikit-learn**
  - Python library for machine learning.
- **Matplotlib and Seaborn**
  - Python libraries for data visualisation.

References

You can also use the following resources for further reading.

1. Introduction to Deep Learning [Link]
2. Introduction to Neural Networks [Link]
3. Deep Learning [Link]
4. What is Deep Learning? [Link]
5. Fundamentals of Deep Learning Activation Functions [Link]